

Article

Evaluation of Urban Vibrancy and Its Relationship with the Economic Landscape: A Case Study of Beijing

Chen Jia ¹ , Yuefeng Liu ¹, Yunyan Du ² , Jianfeng Huang ¹  and Teng Fei ^{3,*} 

¹ Institute of Remote Sensing and Geographic Information System, Beijing Key Lab of Spatial Information Integration and Its Applications, Peking University, Beijing 100871, China; jiachen07@pku.edu.cn (C.J.); yuefengliu@pku.edu.cn (Y.L.); jianfenghuang@pku.edu.cn (J.H.)

² Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Science, Beijing 100864, China; duyuy@lreis.ac.cn

³ School of Resource and Environmental Science, Wuhan University, Wuhan 430072, China

* Correspondence: feiteng@whu.edu.cn

Abstract: As one of the essential indicators for the development of a city, urban vibrancy plays an important role in evaluating the quality of urban areas and guiding urban construction. The development of spatial big data makes it possible to obtain information on user trajectories and the built environment, providing support for the evaluation of urban vibrancy. However, previous studies focused on the number of regional activities when evaluating urban vibrancy and ignored diversity, which was produced by diverse economic landscapes. In this paper, using mobile phone trajectory data, we propose a method for evaluating urban vibrancy from two dimensions: busyness and diversity, based on the improved PageRank algorithm and an index of entropy. Furthermore, in order to explore the relationship between urban vibrancy and the economic landscape, we construct an economic landscape index system based on multi-source data, including points of interest (POIs), roads, building footprints, house prices, the gross domestic product (GDP), and population data. Then, multiple linear regression is utilized to model the relationship between urban vibrancy and the urban economic landscape. The results show that combining busyness and diversity can better characterize urban vibrancy than any single indicator, and the adjusted R-squared (R^2) of the regression with economic landscape reaches 0.59.

Keywords: urban vibrancy; PageRank; entropy; urban landscape characteristics



Citation: Jia, C.; Liu, Y.; Du, Y.; Huang, J.; Fei, T. Evaluation of Urban Vibrancy and Its Relationship with the Economic Landscape: A Case Study of Beijing. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 72. <https://doi.org/10.3390/ijgi10020072>

Academic Editors: Wolfgang Kainz and Giuseppe Borruo

Received: 30 December 2020

Accepted: 9 February 2021

Published: 13 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Urban vibrancy is one of the vital indicators to evaluate urban regional quality and guide urban construction. Jacobs [1] described urban vibrancy by stating that “liveliness and variety attract more liveliness; deadness and monotony repel life”, and she believed that urban vibrancy is highly affected by a dense concentration of people associated with social and economic activities. Lynch [2] considered the influence of the physical properties of a city and considered that vibrancy has three main components: urban morphology, urban function, and urban society. Montgomery [3] pointed out that good cities tend to be a balance of a reasonably ordered and legible city form, and places of many and varied comings and goings, meetings, and transactions. On the one hand, the stimulation of urban vibrancy requires well-planned and fully functional physical facilities; on the other hand, it also requires benign interactions between residents and urban facilities. Therefore, Nicodemus [4] defined neighbourhood vibrancy as the synergy between people, activities, and values in a place, which increases the vibrancy of a community and spurs economic opportunity.

The cause of urban vibrancy is an important research topic. Only by probing its origins can we effectively improve the quantification of urban vibrancy. Research studies on the causes have mostly focused on urban landscapes. Meng and Xing [5] delineated multi-level

characteristics of urban landscapes using multi-source data, including points of interest (POIs), land use, secondary roads, and buildings, and explored the relationship between urban landscape and urban vibrancy. Huang et al. [6] measured the surface attributes to capture the effects of the underlying urban vibrancy from multiple perspectives, utilizing multi-source spatial big data. Wu et al. [7] believed neighbourhood vibrancy can be represented by an index that contains five measures: inter-circulation systems, external traffic system, accessibility, density, and mixed land use.

The first key challenge of urban computing is data acquisition [8]. In previous studies, a large amount of research data on cities has come from questionnaires of urban residents, such as interviews, census, and survey data [9,10]. However, although questionnaires are often used to quantify the attributes of urban space that depend on subjective perception in urban design and implementation, the description of cities is not objective enough. Moreover, the sample sizes have been small and the update speed is slow, which cannot fully reflect the rapid changes in cities. In recent years, with the development of information and communication technologies, especially the popularity of mobile devices, it has become possible to obtain data from individuals and to study their behaviour in cities; such data can include mobile phone positioning data [11,12], social media data [13,14], taxi trajectories [15,16], and public transport smart card records [17,18].

When using data from individuals to evaluate urban vibrancy, most of the previous studies focused more on the quantity, which is the number or density of individual activities in a confined area. Yue et al. [19] defined urban vibrancy as the accumulated number of individuals in a place. Meng and Xing [5] defined urban vibrancy as the sum of individuals' kernel density in an area. Wu et al. [7] defined vibrancy by the percentage of out-of-home non-work activities in a given neighbourhood. However, another dimension of urban vibrancy, diversity, should not be ignored.

Quigley [20] pointed out that urban diversity affects the level of output and the level of well-being achievable in a city. Scholars have done considerable research on diversity in cities, such as the diversity of land use. Van Eck and Koomen [21] assessed the concentration of urbanisation and land-use diversity in a grid-based land-use modelling system. Hao et al. [22] used statistical models to explain land-use diversity, in order to explore urban villagers' land-use pattern. Tan and Wu [23] adopted information entropy as diversity in the regulation of the land use structure. Sulis et al. [24] claimed that urban diversity and urban vibrancy have a certain relationship, and they used mobile card data to calculate mobility patterns of diversity. The important impact of diversity on urban spatial quality and urban vibrancy has been extensively studied and confirmed. However, existing work ignores diversity in the calculation of urban vibrancy. On the other hand, previous studies on diversity have mostly been based on static built environments, such as POI data [19,25], land use, and land cover data [26], which is based on the perspective of design and planning. For highly dynamic cities, there are advantages to describing spatial diversity through the types of human activities from the perspective of citizens.

This study aims to evaluate urban vibrancy from a more comprehensive perspective, adding diversity as the new dimension. Based on trajectory data, diversity is calculated from the perspective of residents' activities, not from the static built environment. In order to verify our method and study the causes of urban vibrancy, we construct an economic landscape index system to explore the relationship between urban vibrancy and economic landscape.

The rest of this article is organized as follows. In Section 2, the research area and data used in this study are introduced. In Section 3, the method of evaluating the urban vibrancy and analysing the relationship between urban vibrancy and the economic landscape is elaborated in detail. Section 4 presents the results of our case study. In Section 5, the results and the limitations of our research are discussed. In the last part, brief conclusions are provided.

2. Study Area and Data Material

2.1. Study Area

As shown in Figure 1, the central region of Beijing is selected as the study area. Beijing is the capital of China and also the political and cultural centre of China. There are 16 districts in Beijing, with a total area of 16,410.54 square kilometres, a built-up area of 1485 square kilometres, and a permanent population of 21.5 million. Beijing has a typical concentric circle structure, and most of the social and economic activities are gathered inside the fifth ring road. Therefore, we selected the area within the fifth ring road as the research area.

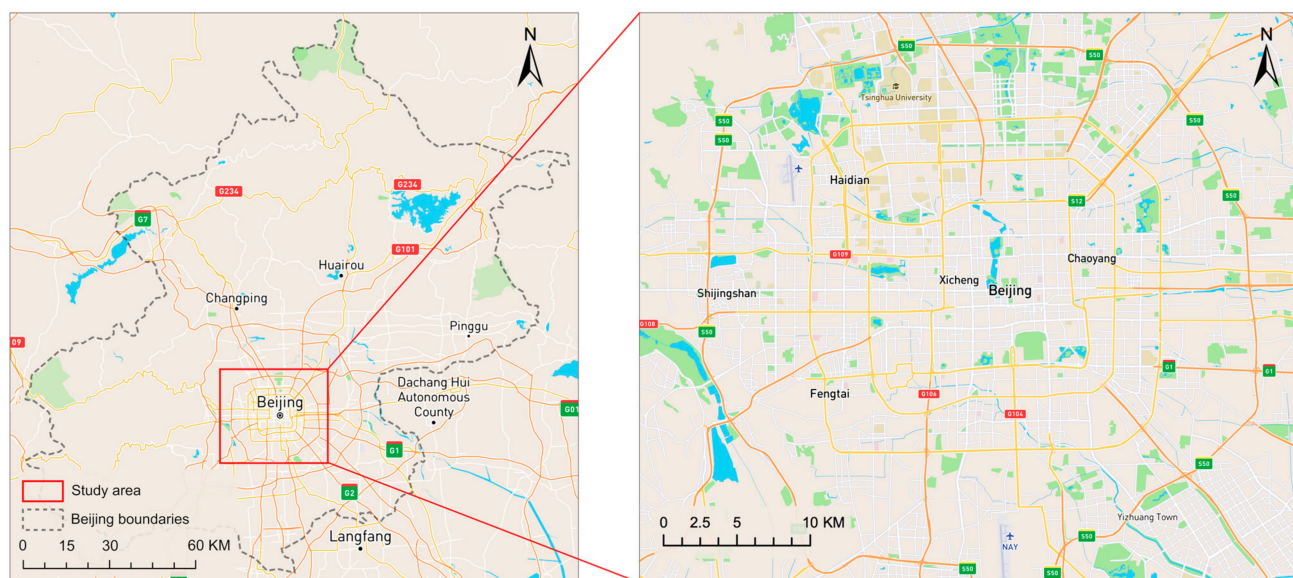


Figure 1. Study area of Beijing, China. (External source: Mapbox, Inc.).

The selection of research units is important [13]. Thiessen polygons [27], grids [6,13], and traffic analysis zones (TAZs) [28] are commonly used. Considering the homogeneity and regularity of the grid, we adopted a grid of 1 km * 1 km as the research unit, of which there was a total of 1840.

2.2. Data and Preprocessing

2.2.1. Human Mobility Data

Mobile phone trajectory data can represent human mobility. Our study data are collected from the Beijing mobile phone communication service provider and contain a total of 1,200,000 random users' trajectories on 27 December 2016, which represents 5.5% of Beijing's permanent residential population. Each trajectory consists of chronological GPS positioning points. Every time a user makes or receives a call, sends or receives a short message, and connects to the network, a GPS positioning record is generated.

Stay point detection is one of the crucial steps in the preprocessing of mobile phone trajectory data. As the mobile phone tracking data contain much redundant information, stay point detection can help find places that really attract people and generate vibrancy.

A stay point represents a geographic region where a user stays over a certain time interval [29]. Let $T = \{p_1, p_2, \dots, p_n\}$ be a trajectory composed of chronological GPS positioning points. Each point is represented as $p = (x, y, t)$, where x, y are geographic locations and t is the time stamp. A stay point s can be represented as a group of consecutive GPS positioning points $P = \{p_m, p_{m+1}, \dots, p_{m+n}\}$, where the distance between every two points is less than a distance threshold D_{th} , and the time interval is greater than the time

interval T_{th} . Then, s can be represented by the mean centre of this group of points, start time, and end time, as $s = (x, y, t_{start}, t_{end})$, where

$$x = \frac{\sum_{i=0}^n p_i \cdot x}{n} \quad (1)$$

$$y = \frac{\sum_{i=0}^n p_i \cdot y}{n} \quad (2)$$

$$t_{start} = p_m \cdot t \quad (3)$$

$$t_{end} = p_{m+n} \cdot t \quad (4)$$

In this study, considering the average distance between base stations and walking speed, the distance threshold D_{th} is set to 500 m, and the time interval T_{th} is 0.5 h.

Stay point detection is performed on the mobile phone trajectory data. Table 1 shows the basic statistics of the stay point detection results. The methods presented later are based on extracted stay points.

Table 1. Basic statistics of stay point detection results.

Number of Trajectories	Average Length of Trajectories	Number of Positioning Points per User	Number Of Stay Points per User	Percentage of Points after Detection
1,200,000	24.74 km	31.23	1.49	4.77%

2.2.2. Economic Landscape Data

1. Urban built environment data.

A total of four data sources were used to represent the urban built environment, i.e., the POI data, road network data, building footprint data, and house price data, as shown in Figure 2. Among them, POI data, road network data, and building footprint data for 2016 were derived from Amap (<https://www.amap.com> (accessed on 30 December 2020)), a digital map content, navigation, and location service solution provider in China.

The POI data in the study area contain 103,576 points in total, 15 major categories, and 48 subcategories, as shown in Figure 2a. Each POI is composed of its spatial location and category code. The total length of the road network data is 8,860,750 km, as shown in Figure 2b. Building footprint data contain 297,211 polygons, including their shapes, locations, and floors, covering a total area of 283.57 km², as shown in Figure 2c. All the roads and building polygons are segmented and identified according to the grids.

House price data were obtained from LianJia (<https://www.lianjia.com> (accessed on 30 December 2020)), one of China's largest real estate service platforms. There are 301,633 housing transaction records from 2012 to 2016. Due to the lack of house price data in some areas, the inverse distance weighting (IDW) method is used for spatial interpolation, where the power value is set to 2. The spatial distribution of the house price is shown in Figure 2d.

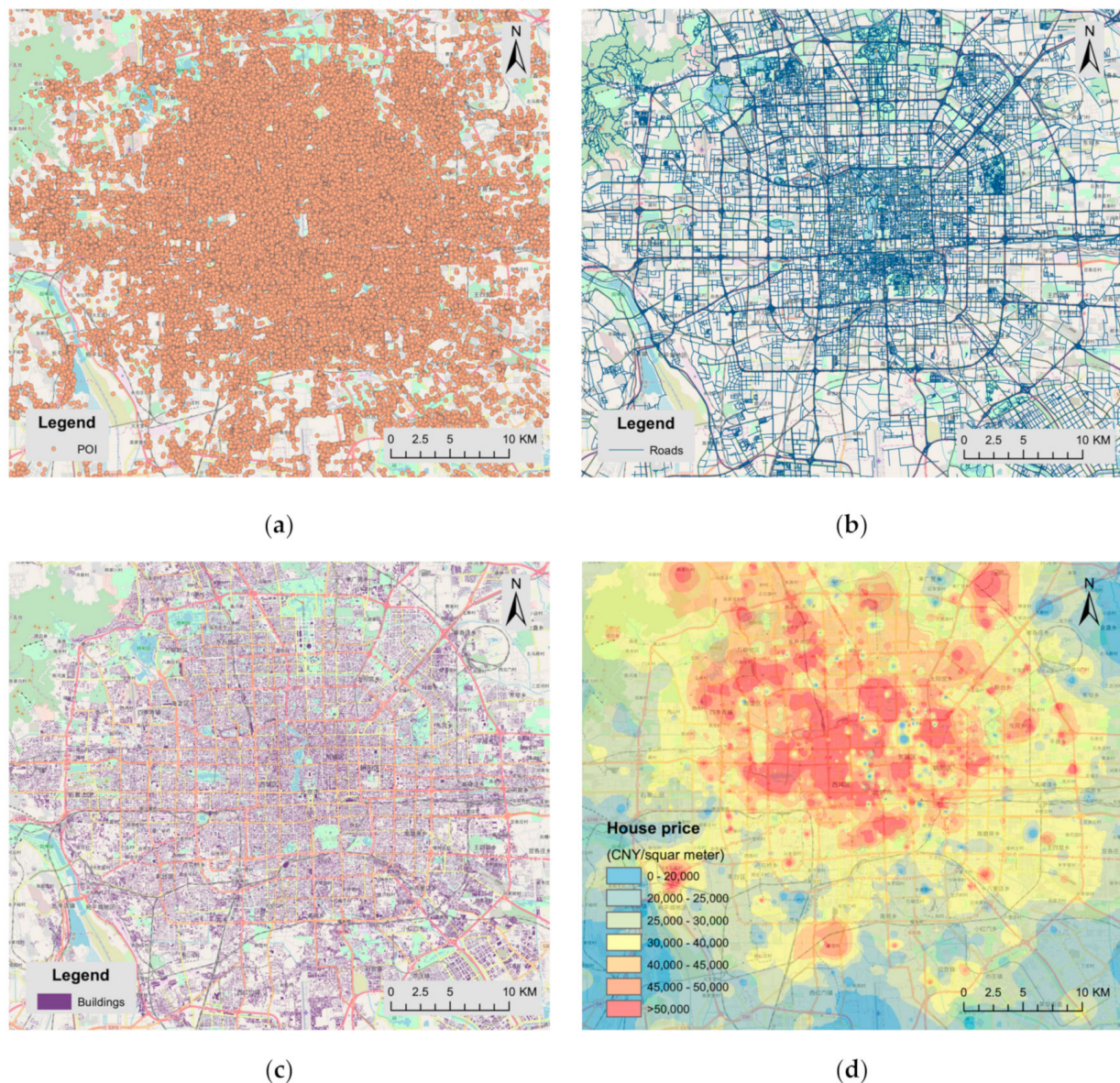


Figure 2. Urban built environmental data. (a) POI data; (b) Road network data; (c) Building footprint data; (d) House price data.

2. Gross domestic product (GDP) and population data.

The spatial distribution data of GDP and population were provided by the Resource and Environmental Science Data Centre of the Chinese Academy of Sciences (<http://www.resdc.cn> (accessed on 30 December 2020)). These two data sets reflect the detailed spatial distribution of China's GDP and population in 2015. Both data sets contain raster data, and each raster represents the GDP and the total population within the grid area (1 square kilometre) [30,31]. Based on the statistical data of administrative districts, these two data sets comprehensively consider the types of land use, night light intensity, and residential density. The multi-factor weight distribution method is used to allocate GDP and population to grids.

3. Methodology

The overall research framework is shown in Figure 3. Urban spatial vibrancy in this study is described in two dimensions: busyness and diversity. The improved PageRank algorithm (IPA) is used to evaluate the busyness of a place in Section 3.1, and diversity is

described in terms of entropy in information theory in Section 3.2. In order to verify the urban vibrancy evaluation method and further explain the causes of urban vibrancy, we perform a regression analysis between the two dimensions of urban vibrancy and economic landscape, which is introduced in Section 3.3.

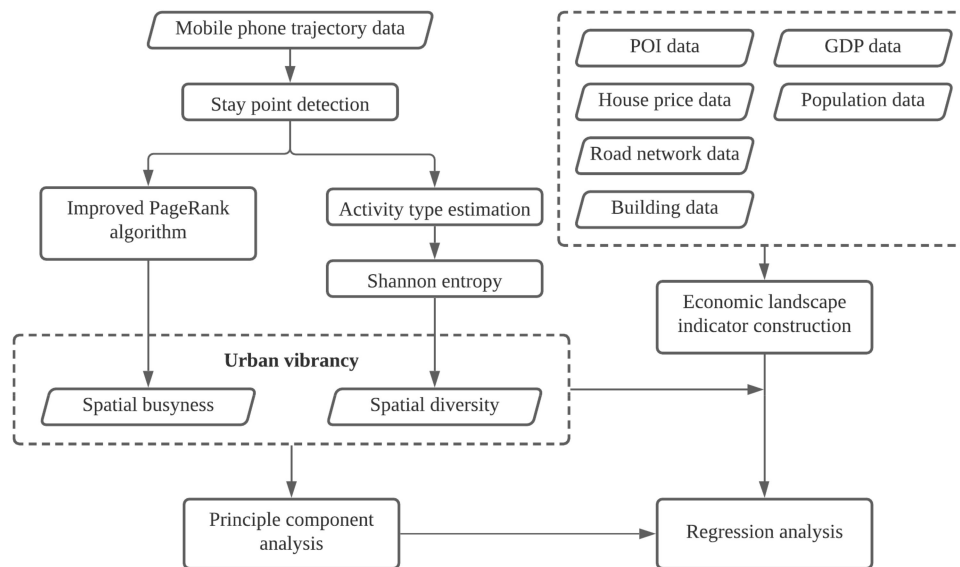


Figure 3. Research framework.

3.1. Evaluation of Spatial Busyness with the Improved PageRank Algorithm

Spatial busyness is a quantitative measure of the intensity of urban activity. In our previous work, we proposed an improved PageRank algorithm, which extends the traditional PageRank method to a weighted bipartite graph [32]. Different from the conventional methods which measure spatial busyness by the number of people, our method takes into account the individual differences in the contributions to spatial busyness. It is assumed that users who travel more widely and more intensively carry more weight with respect to the ability to “splash” vibrancy in a given neighbourhood. In detail, we consider that busier users, who have visited more places, have a higher weight in the contribution of the busyness, and less busy users have a lower weight.

The algorithm involves two types of entities: user and area. When a user visits an area, he or she will “vote” for this area, and the area also “votes” for the user. Then, the busyness of a user is equal to the sum of all votes cast by areas he or she has visited, and the busyness of an area is equal to the sum of all votes cast by the visited users. After a finite number of iterations, the busyness of users and areas converge to a stable value. Finally, spatial busyness can be obtained from the busyness of each area.

Let u be a user and a be an area. Then, S_u is the set of areas user u has visited, and T_a is the set of users who have visited area a . $|S_u|$ and $|T_a|$ are the numbers of elements in the two sets. The busyness R_u of user u and the busyness R_a of area a are defined as follows:

$$R_u = \frac{1-d}{|S_u|} + d \sum_{x \in S_u} \frac{R_x}{|T_x|} \quad (5)$$

$$R_a = \frac{1-d}{|T_a|} + d \sum_{x \in T_a} \frac{R_x}{|S_x|} \quad (6)$$

Parameter d is the damping factor, following relevant studies [33], which is usually set to 0.85. The source code of the improved PageRank algorithm is provided in Appendix A.

3.2. Evaluation of Spatial Diversity with Entropy

Spatial diversity is a measure of activity heterogeneity. Our quantification of the diversity of each spatial grid is based on the diversity of the types of activities of the population in this area. We discriminate the types of users' activities based on the stay points extracted from trajectory data and POI category data.

Firstly, in order to describe the purpose of users' activities, we reclassify POI categories according to daily behaviour types, which are divided into 15 categories: catering, accommodation, retail, automotive services, finance, education and culture, medical and health, sports and leisure, public facilities, commercial services, residential services, companies, transportation, scientific research institutions, and agriculture. Then, we assign reasonable opening hours to subcategories.

Secondly, Gaussian kernel density estimation is used to calculate the probability of each POI near the stay point becoming a destination, as follows:

$$y = f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (7)$$

where y is the probability density at which the POI becomes a destination; x is the distance from this POI to the stay point; σ is the standard deviation, which is set at 500 m, and μ is the mean value, which is 0.

Based on the distance between the base stations, the search radius is set to 1 km. The probability that all the opening POIs within the search radius of the stay point become the destination is calculated. The activity type of the POI with the highest probability is taken as the activity type of the stay point.

Shannon entropy [34] is the most commonly used indicator when measuring the degree of system diversity. Similarly, in the complex urban system composed of people and land, we also selected Shannon entropy to measure spatial diversity.

Given an area where the activity types of all stay points are in the set $\{c_1, c_2, \dots, c_n\}$, c_i is the i -th activity type of stay point, and $|c_i|$ is the number of stay points of type $|c_i|$. The probability distribution is:

$$P(X = c_i) = p_i = \frac{|c_i|}{\sum_{i=1}^n |c_i|}, \quad i = 1, 2, \dots, n. \quad (8)$$

The entropy of the area is calculated as:

$$H(X) = - \sum_{i=1}^n p_i \log p_i \quad (9)$$

$$0 \leq H(X) \leq \log n \quad (10)$$

Finally, spatial diversity can be obtained from the entropy of each area. The larger the $H(x)$, the higher the degree of diversity.

3.3. Analysis of the Relationship between Urban Vibrancy and Economic Landscape

Previous studies have shown that factors affecting urban vibrancy include population density, building density, mixed land use, and accessibility [3,35,36]. In addition, Lu et al. [37] found that house prices have a strong positive correlation with urban vibrancy, since house prices are directly related to spatial location. Places with higher house prices have more convenient transportation, parks, and larger commercial districts. Economic vibrancy is also an important part of urban vibrancy [38,39], and GDP is the important measurement index of urban economic capacity [40]. To explore the causes of urban vibrancy, we construct multi-level urban economic landscape indicators, which include three parts, the built environment factor, GDP factor, and population factor. In the built environment factor, the density of POIs represents the density of urban functions, the total length of roads indicates road features, and the house price reflects the land value. We also use the

density, gross floor, height, and landscape shape index (BLSI) [41] of buildings for depicting building characteristics. Table 2 shows the specific definitions of each indicator.

Table 2. Definition of the urban economic landscape.

Factor	Indicators	Definition
Built environment factor	Density of POIs	Number of POIs divided by the area.
	Total length of roads	Roads comprise five levels: expressways, first-level highways, second-level highways, third-level highways, and fourth-level highways.
	Average house price	Average housing transaction price, reflecting the land value of an area.
	Standard deviation of house price	Standard deviation of housing transaction prices, reflecting regional land value differences.
	Building density	Total base area of all buildings divided by the total land area, which can reflect the vacancy rate and building density.
	Gross floor area	Sum of the total floor area in an area.
	Standard deviation of gross floor area	Standard deviation of the total floor area in an area.
	Average building height	Total height of buildings divided by the building numbers.
	Standard deviation of building height	Standard deviation of building height in an area.
	Landscape shape index of building (BLSI)	$\frac{\sum_i E_i}{2\sqrt{\pi \sum_i A_i}}$, where E_i is the circumference of the i th building and A_i is the area of the i th building.
GDP factor	Gross regional product	Economic output in a specific area, obtained by resampling the GDP kilometre grid data sets.
Population factor	Population density	Number of individuals divided by the area, obtained by resampling the population kilometre grid data sets.

Multiple linear regression is utilized to model the relationship between urban vibrancy and urban economic landscape. The independent variables are the urban economic landscape indicators, and the dependent variables are spatial busyness and spatial diversity. The selection of indicators uses stepwise regression.

In order to compare the effects of the area busyness and area diversity separately and their combined effects in terms of the predictivity with respect to the economic landscape, principal component analysis (PCA) was performed on busyness and diversity, and the principal component with the largest eigenvalue was extracted and used as a dependent variable. The goodness of fit of the three models (predictive variable with busyness, diversity and the first principal component of the two) was evaluated.

4. Results

4.1. Urban Vibrancy Results in Beijing

Figure 4 shows the spatial distribution of urban vibrancy. Figure 4a shows the urban busyness, while Figure 4b shows the urban diversity, both of which are normalized to between 0 and 1. Although the busyness and diversity as a whole show a decreasing trend from the centre to the surroundings, in line with the circle structure of Beijing, there are still significant differences between the two.

According to Figure 4a, the busiest area is not in the centre of Beijing. The area between the Second Ring Road and Fourth Ring Road is busy. The south is busier than the north, and there are two high-busyness centres in the south-west and south-east areas. The spatial distribution of diversity is entirely different from that of busyness, as shown in Figure 4b. The area inside the Fourth Ring Road shows high diversity, which also means that the urban facilities inside the Fourth Ring Road are relatively complete, and the degree of mixed land use is relatively high. Outside the Fourth Ring Road, diversity gradually decreases, and areas with high diversity are mostly located along traffic lines.

Figure 5 shows the distribution of busyness and diversity of all grids in two-dimensional coordinates. To further explore the difference between busyness and diversity and their

spatial distribution, we assign the “H” label to the grids ranked in the top 30% of busyness, and the “L” label to the grids ranked in the bottom 30%, and the same applies to diversity. Then, we obtain four types of grids, namely, “HH”, “LL”, “HL”, and “LH”, where “HL” means high busyness and low diversity, as shown in Table 3. Figure 6 shows the spatial distribution of these four types of grids.

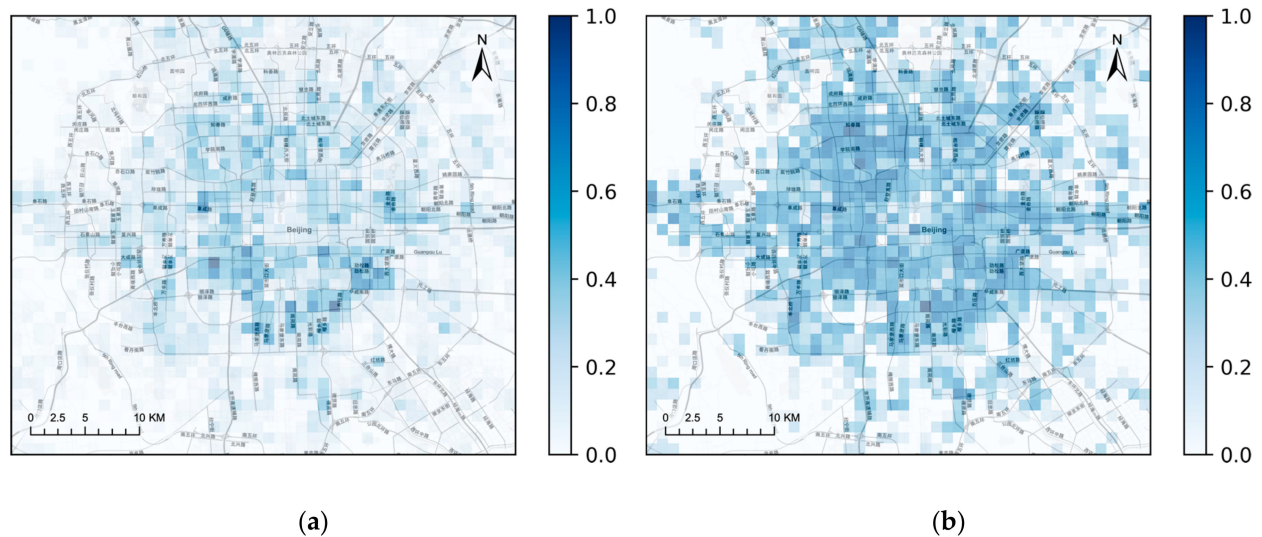


Figure 4. Urban vibrancy spatial distribution in Beijing. (a) The urban busyness; (b) The urban diversity.

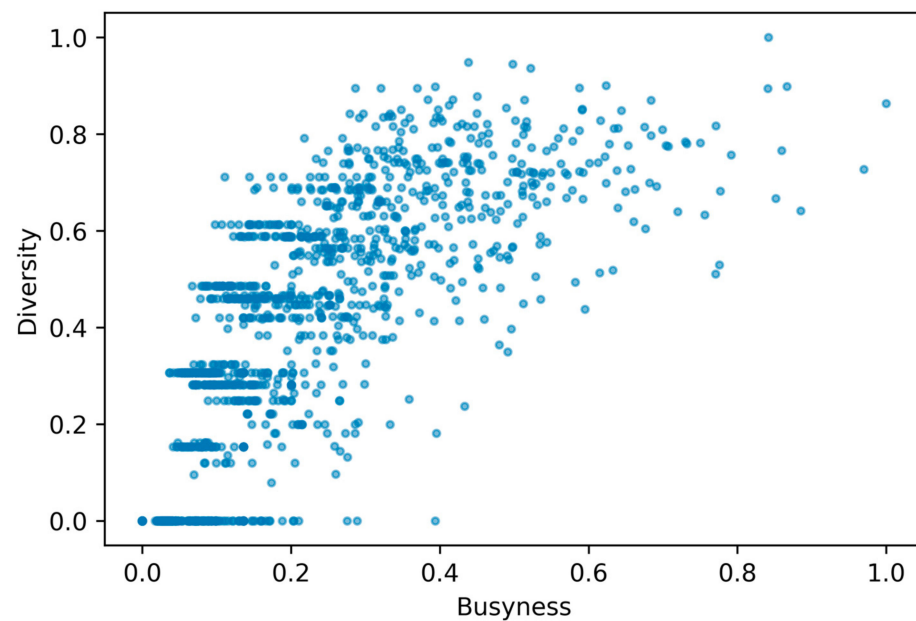


Figure 5. Busyness and diversity.

Table 3. Descriptions of the four grid categories.

Categories	Busyness Label	Diversity Label	Descriptions
HH	H	H	Busyness ranks in the top 30%, and diversity ranks in the top 30%.
LL	L	L	Busyness ranks in the bottom 30%, and diversity ranks in the bottom 30%.
HL	H	L	Busyness ranks in the top 30%, and diversity ranks in the bottom 30%.
LH	L	H	Busyness ranks in the bottom 30%, and diversity ranks in the top 30%.

It can be seen that HH and LL are the most numerous categories. HH is mainly concentrated in the area inside the third ring road, and LL is mainly distributed in the area outside the fifth ring road. The urban centre is more prosperous, with higher busyness and diversity, while the outer suburbs are less busy and diverse. However, in addition to these two categories, we also see a small number of regions with HL and LH patterns. HL means high busyness and low diversity. This type of area is usually a single functional area and requires supporting facilities in the surrounding area to meet daily work and life needs, so there will be more interactions with surroundings. Because its dominant function is well developed, it is also attractive to more distant residents. Typical HL areas include industrial parks, large residential areas, etc. LH means low busyness and high diversity. Its facilities and services are relatively complete, which can sufficiently cover daily work and life needs. However, the number of people it serves is relatively small—mainly local residents. Its functional facilities also have no comparative advantage, so it is not attractive to residents in more distant areas. Typical LH areas include urban suburbs and so on, which are self-sufficient undeveloped communities. In short, busyness and diversity reflect the different characteristics of urban areas.

**Figure 6.** Cont.



Figure 6. Spatial distribution of the grids of four categories. (a) HH; (b) LL; (c) HL; (d) LH.

4.2. The Relationship between Urban Vibrancy and the Economic Landscape

The stepwise method is used for multiple linear regression. Tables 4 and 5 display the regression model results between spatial busyness and economic landscape (Table 4), and between spatial diversity and economic landscape (Table 5).

When modelling the relationship between spatial busyness and economic landscape, seven variables are selected. The R^2 of the model reaches 0.55. Similarly, when modelling the relationship between spatial diversity and economic landscape, the R^2 of the model reaches 0.53. At the significance level of 0.05, all variables are significant.

On the other hand, we take the first principal component of busyness and diversity, as shown in Table 6, and perform multiple linear regression of economic landscape and component scores, with R^2 reaching 0.59, as shown in Table 7. Compared with the linear regression result of any single indicator (busyness or diversity) and the economic landscape, this result yields a higher goodness of fit. It can be inferred that either busyness or diversity alone cannot fully reflect the economic landscape. Combining these two indicators can better characterize urban vibrancy than any single indicator.

To eliminate multicollinearity, variables with variance inflation factor (VIF) values greater than 10 are removed. The above regression results have passed the residual normality and variance homogeneity tests.

Table 4. Regression model results between spatial busyness and economic landscape.

Variables	B	Std. Error	Beta	VIF
(Constant)	0.01	0.01		
Gross floor area	1.38×10^{-7}	9.08×10^{-9}	0.48 ***	4.09
Population density	7.32×10^{-6}	6.52×10^{-7}	0.25	1.99
Standard deviation of the gross floor area	-6.90×10^{-6}	7.26×10^{-7}	-0.24	2.62
Density of POIs	3.75×10^{-4}	5.18×10^{-5}	0.17	2.37
Standard deviation of the building height	0.02	3.26×10^{-3}	0.34	9.86
Average building height	1.62×10^{-3}	3.05×10^{-4}	-0.22	6.99
Average house price	-6.76×10^{-7}	3.08×10^{-7}	-0.04 *	1.57

$N = 1840$, $R = 0.75$, $R^2 = 0.55$, Adjusted $R^2 = 0.55$, $F = 325.69$; * $p < 0.05$, *** $p < 0.001$.

Table 5. Regression model results between spatial diversity and economic landscape.

Variables	B	Std. Error	Beta	VIF
(Constant)	−0.05	0.02		
Gross floor area	2.10×10^{-7}	1.50×10^{-8}	0.43 ***	3.73
Population density	1.89×10^{-5}	2.27×10^{-6}	0.38	8.15
Density of POIs	6.96×10^{-4}	8.98×10^{-5}	0.19	2.40
Standard deviation of the gross floor area	$−6.51 \times 10^{-6}$	1.10×10^{-6}	−0.14	2.03
Gross regional product	$−4.56 \times 10^{-7}$	8.35×10^{-8}	−0.23	7.10
Average house price	7.38×10^{-7}	1.70×10^{-7}	0.10	2.23
Landscape shape index of buildings (BLSI)	0.03	0.01	0.05 *	1.31

$N = 1840$, $R = 0.73$, $R^2 = 0.53$, Adjusted $R^2 = 0.53$, $F = 294.90$; * $p < 0.05$, *** $p < 0.001$.

Table 6. Factor analysis result.

Two Dimensions of Urban Vibrancy	Descriptive Statistics			Factor Loading 1	Communalities	
	N	Mean	Std. Deviation		Initial	Extraction
Busyness	1840	0.16	0.17	0.95	1.00	0.91
Diversity	1840	0.29	0.28	0.95	1.00	0.91

Table 7. Regression model result between component scores and economic landscape.

Variables	B	Std. Error	Beta	VIF
(Constant)	−1.13	0.04		
Gross floor area	8.53×10^{-7}	5.17×10^{-8}	0.50 ***	4.10
Population density	6.34×10^{-5}	7.26×10^{-6}	0.36	7.66
Standard deviation of the gross floor area	$−3.44 \times 10^{-5}$	4.01×10^{-6}	−0.20	2.48
Density of POIs	2.21×10^{-3}	3.03×10^{-4}	0.17	2.51
Gross regional product	$−1.21 \times 10^{-6}$	2.70×10^{-7}	−0.18	6.80
Average house price	1.77×10^{-6}	5.73×10^{-7}	0.07 **	2.32
Total length of roads	1.23×10^{-5}	5.56×10^{-6}	0.04 *	1.69
Average building height	2.05×10^{-2}	9.67×10^{-3}	0.05	2.34

$N = 1840$, $R = 0.77$, $R^2 = 0.59$, Adjusted $R^2 = 0.59$, $F = 328.92$; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

5. Discussion

5.1. Evaluating Urban Vibrancy in Busyness and Diversity

In Section 4.1, we measured urban vibrancy, and identified four grid patterns based on busyness and diversity, naming them as HH, LL, HL, and LH. Figure 7 and Table 8 present examples for more detailed analysis.

An example of HH is located within the Second Ring Road, including Beijing Railway Station, residential communities, department stores, and middle schools. Beijing Railway Station has a large passenger flow, transporting millions of passengers every year. In addition to the railway station, the residential, commercial, educational, and other facilities are complete, so the busyness and diversity are high.

An example of LL is located in Guogongzhuang Village, Fengtai District. It is mainly farmland and agricultural facilities, with a single function and a small flow of people. Therefore, busyness and diversity are low.

An example of HL is between the Fourth and Fifth Ring Roads, including the Olympic Sports Centre and Olympic Forest Park. The Olympic Sports Centre has hosted a number of sports events, including the 2008 Summer Olympics. It is a sports centre that integrates competition, training, and fitness. Due to the complete sports facilities in the area, it has a strong attraction to the surrounding area and has high busyness. However, because of its single function, it has low diversity.

An example of LH is located in the north-western suburbs, including Fragrant Hills and Beijing Botanical Garden. Fragrant Hills is a natural scenic spot and public park.

Because it is far from the city centre, it serves fewer residents and has low busyness. However, relying on the tourism industry, it has developed various supporting facilities and formed an independent community, thus having a high diversity.

The quantitative measurement of urban vibrancy is a difficult problem. In previous studies, urban vibrancy has often been described in a single dimension, mostly the degree of busyness. However, in our research, we found that diversity is indeed an indispensable part of portraying urban vibrancy, and it is not exactly the same as busyness. From Figure 4, we can see that busyness and diversity present different forms and trends in space. It is more comprehensive and accurate to use two dimensions to represent urban vibrancy. Through the combination of the two dimensions, we can also distinguish regional vibrancy patterns to a certain extent.

In the measurement of spatial busyness, we use the improved PageRank method, which takes into account the differences in personal activities: high-busyness users contribute more to spatial busyness, and further improve the city's busyness, and the busyness can be better distinguished. In the measurement of diversity, we use Shannon entropy. This method first distinguishes the types of human activities and then calculates the diversity of activities in the region. In addition to Shannon entropy, the calculation of diversity can also consider richness, the Simpson index, etc. From the perspective of residents, this research explores the diversity of their daily activities. In future research, the diversity of land use can also be used from the perspective of urban functions.

Although this study proposes a method to measure urban vibrancy from two dimensions of busyness and diversity based on spatial big data, we should always be aware of the biases of these data sources, such as sampling bias. In addition, due to the limitation of data acquisition, we only studied the day of 26 December 2016 which is a weekday. However, weekends or holidays, etc., may be different, and unexpected events, weather and other conditions will also cause changes in vibrancy. Future work includes exploring the urban vibrancy over a longer period and the impact of different events on vibrancy.

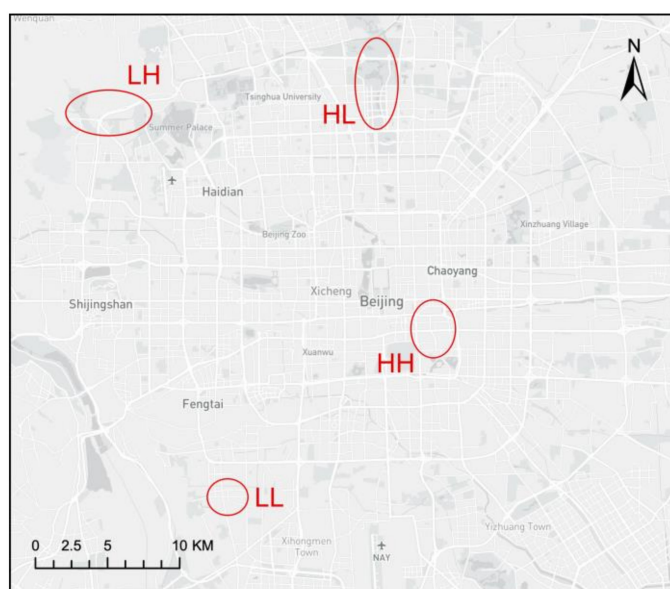


Figure 7. Examples of four patterns: HH, LL, HL, and LH.

Table 8. Geospatial information and urban vibrancy of the example area.

Pattern	Busyness	Diversity	Location and Facilities
HH	0.87	0.90	Beijing Railway Station, Xijingjiayuan Community, New World Department Store, Beijing No. 11 Middle School
LL	0.05	0.06	Guogongzhuang Village
HL	0.48	0.18	National Olympic Sports Centre, Olympic Forest Park
LH	0.09	0.46	Fragrant Hills, Beijing Botanical Garden

5.2. Exploring the Relationship between Urban Vibrancy and Economic Landscape

Urban vibrancy is the result of the interaction between human activities and well-planned urban facilities. We put forward an economic landscape index system, and separately model and analyse the relationship between the economic landscape and the two dimensions of urban vibrancy, spatial busyness and spatial diversity. The results show that there is a correlation between these two dimensions of urban vibrancy and the economic landscape, and two dimensions can better represent urban vibrancy than any single dimension.

At the same time, different economic landscapes illustrate different interpretations of the two dimensions of urban vibrancy. For spatial busyness, gross floor area, building height, and population density contribute more, while for spatial diversity, gross floor area, population density, and gross regional product contribute more.

It is worth noting that gross regional product is negatively correlated with diversity. Areas with high gross regional product often have a single function, so their diversity is low. In addition, house prices are negatively correlated with busyness and positively correlated with diversity. This may be due to the fact that highly prosperous areas are usually noisy and uninhabitable, hence the lower house prices. Meanwhile, higher priced residential areas tend to be equipped with an abundance of amenities, so the spatial diversity is higher.

However, there are a great many factors influencing urban vibrancy, and we only selectively discussed some common factors due to the restriction of data, mainly focusing on the built environment, GDP and population. Other factors, such as neighborhood density, land use mix and green space, which have been reported to have important impact on urban vibrancy, were not included. Besides, the interesting relationship between social cohesion and urban vibrancy was also not explored—they share many common factors, but the mechanisms of how these factors affect social cohesion and urban vibrancy still need further research [42].

Overall, through the analysis of the economic landscape and urban vibrancy, on the one hand, we confirmed the effectiveness of the proposed urban vibrancy algorithm and found the difference between its components of busyness and diversity. On the other hand, this study also revealed some of the causes of urban vibrancy, which is essential for urban planning and construction.

6. Conclusions

Urban vibrancy is a comprehensive aspect that has been used to describe and evaluate the development of a city. This paper proposes a method to quantitatively evaluate urban vibrancy from mobile phone-derived trajectory data. This method involves the calculation of two vibrancy indices, namely, busyness and diversity.

We have employed an improved PageRank algorithm for the calculation of area busyness. In this process, every person who was active in an area “votes” for the area’s vibrancy, and by design, people do not have the same voting weight: the area’s busyness is more affected by the number of people who had been to more areas in the city over a given period of time. On the other hand, we calculated the diversity of areas as well. In this process, diversity is not only defined by the types of POIs in an area, but more importantly, it is determined by the diversity of the purpose of people’s stay in the area. This diversity index is obtained by Gaussian kernel density estimation and is represented

by Shannon entropy. Based on above indices proposed, we have identified four types of vibrancy patterns in urban space of Beijing.

In order to verify whether the urban vibrancy, represented by area busyness and area diversity, is more explanatory of the economic activity of a city than using either one indicator, we used multiple sources of data to characterize the economic landscape of a sample city. The result shows that a model combining busyness and diversity can explain the economic landscape of the city better than using one single factor.

In the context of rapid urban development, vibrancy is an important key for cities to attract talents, investment, and improve competitiveness. The method proposed in this article presents a scientific, objective yet spatially explicit approach of calculating urban vibrancy, which is urgently needed, from mobile trajectory data. It provides a comprehensive reference for future urban planning, spatial evaluation, and site selection.

Author Contributions: Teng Fei and Chen Jia conceived the presented idea and completed the manuscript. Yunyan Du helped with data acquisition and processing. Chen Jia conducted the experiments. Yuefeng Liu and Jianfeng Huang helped improve the experiment and modify the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, Grant No. U1433102.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Source code of the improved PageRank algorithm: <https://github.com/JiaChen-07/UrbanVibrancy> (accessed on 30 December 2020).

References

- Jacobs, J. *The Death and Life of Great American Cities*; Vintage Books: New York, NY, USA, 1961.
- Lynch, K. *Good City Form*; MIT Press: Cambridge, MA, USA, 1984.
- Montgomery, J. Making a city: Urbanity, vitality and urban design. *J. Urban Des.* **1998**, *3*, 93–116. [\[CrossRef\]](#)
- Nicodemus, A.G. Fuzzy vibrancy: Creative placemaking as ascendant US cultural policy. *Cult. Trends* **2013**, *22*, 213–222. [\[CrossRef\]](#)
- Meng, Y.; Xing, H. Exploring the relationship between landscape characteristics and urban vibrancy: A case study using morphology and review data. *Cities* **2019**, *95*, 102389. [\[CrossRef\]](#)
- Huang, B.; Zhou, Y.; Li, Z.; Song, Y.; Cai, J.; Tu, W. Evaluating and characterizing urban vibrancy using spatial big data: Shanghai as a case study. *Environ. Plan. B Urban Anal. City Sci.* **2019**, *47*, 1543–1559. [\[CrossRef\]](#)
- Wu, J.; Ta, N.; Song, Y.; Lin, J.; Chai, Y. Urban form breeds neighborhood vibrancy: A case study using a GPS-based activity survey in suburban Beijing. *Cities* **2018**, *74*, 100–108. [\[CrossRef\]](#)
- Zheng, Y.; Capra, L.; Wolfson, O.; Yang, H. Urban computing: Concepts, methodologies, and applications. *ACM Transac. Intell. Syst. Technol.* **2014**, *5*, 1–55. [\[CrossRef\]](#)
- Azmi, D.I.; Karim, H.A. Implications of walkability towards promoting sustainable urban neighbourhood. *Procedia Soc. Behav. Sci.* **2012**, *50*, 204–213. [\[CrossRef\]](#)
- Holian, M.J.; Kahn, M.E. *The Impact of Center City Economic and Cultural Vibrancy on Greenhouse Gas Emissions from Transportation*; Mineta Transportation Institute: San Jose, CA, USA, 2012.
- Gonzalez, M.C.; Hidalgo, C.A.; Barabasi, A.-L. Understanding individual human mobility patterns. *Nature* **2008**, *453*, 779–782. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kang, C.; Liu, Y.; Ma, X.; Wu, L. Towards estimating urban population distributions from mobile call data. *J. Urban Technol.* **2012**, *19*, 3–21. [\[CrossRef\]](#)
- Wu, C.; Ye, X.; Ren, F.; Du, Q. Check-in behaviour and spatio-temporal vibrancy: An exploratory analysis in Shenzhen, China. *Cities* **2018**, *77*, 104–116. [\[CrossRef\]](#)
- Agryzkov, T.; Martí, P.; Tortosa, L.; Vicent, J.F. Measuring urban activities using Foursquare data and network analysis: A case study of Murcia (Spain). *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 100–121. [\[CrossRef\]](#)
- Liu, Y.; Kang, C.; Gao, S.; Xiao, Y.; Tian, Y. Understanding intra-urban trip patterns from taxi trajectory data. *J. Geogr. Syst.* **2012**, *14*, 463–483. [\[CrossRef\]](#)
- Zhang, W.; Li, S.; Pan, G. Mining the semantics of origin-destination flows using taxi traces. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 5–8 September 2012; Association for Computing Machinery: Pittsburgh, PA, USA, 2012; pp. 943–949.

17. Long, Y.; Zhang, Y.; Cui, C. Identifying commuting pattern of Beijing using bus smart card data. *Acta Geogr. Sin.* **2012**, *67*, 1339–1352.
18. Zhong, C.; Huang, X.; Arisona, S.M.; Schmitt, G.; Batty, M. Inferring building functions from a probabilistic model using public transportation data. *Comput. Environ. Urban Syst.* **2014**, *48*, 124–137. [\[CrossRef\]](#)
19. Yue, Y.; Zhuang, Y.; Yeh, A.G.; Xie, J.-Y.; Ma, C.-L.; Li, Q.-Q. Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 658–675. [\[CrossRef\]](#)
20. Quigley, J.M. Urban diversity and economic growth. *J. Econ. Perspect.* **1998**, *12*, 127–138. [\[CrossRef\]](#)
21. Van Eck, J.R.; Koomen, E. Characterising urban concentration and land-use diversity in simulations of future land use. *Ann. Reg. Sci.* **2008**, *42*, 123–140. [\[CrossRef\]](#)
22. Hao, P.; Geertman, S.; Hooimeijer, P.; Sliuzas, R. The land-use diversity in urban villages in Shenzhen. *Environ. Plan. A Econ. Space* **2012**, *44*, 2742–2764. [\[CrossRef\]](#)
23. Tan, Y.; Wu, C.-f. The laws of the information entropy values of land use composition. *J. Nat. Resour.* **2003**, *18*, 112–117.
24. Sulis, P.; Manley, E.; Zhong, C.; Batty, M. Using mobility data as proxy for measuring urban vitality. *J. Spat. Inf. Sci.* **2018**, *2018*, 137–162. [\[CrossRef\]](#)
25. Liu, W.; Wu, W.; Thakuriah, P.; Wang, J. The geography of human activity and land use: A big data approach. *Cities* **2020**, *97*, 102523. [\[CrossRef\]](#)
26. Sapena, M.; Ruiz, L.A. Analysis of urban development by means of multi-temporal fragmentation metrics from LULC data. *ISPRS-Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* **2015**, *40*, 1411. [\[CrossRef\]](#)
27. Li, Y.; Fei, T.; Zhang, F. A regionalization method for clustering and partitioning based on trajectories from NLP perspective. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 2385–2405. [\[CrossRef\]](#)
28. Yao, Y.; Li, X.; Liu, X.; Liu, P.; Liang, Z.; Zhang, J.; Mai, K. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 825–848. [\[CrossRef\]](#)
29. Zheng, Y.; Zhang, L.; Xie, X.; Ma, W.-Y. Mining interesting locations and travel sequences from GPS trajectories. In Proceedings of the 18th International Conference on World Wide Web, Madrid, Spain, 20–24 April 2009; Association for Computing Machinery: Madrid, Spain, 2009; pp. 791–800.
30. Xu, X. China Population Spatial Distribution Kilometer Grid Dataset. Data Registration and Publishing System of Resource and Environmental Science Data Center of the Chinese Academy of Sciences. 2017. Available online: <http://www.resdc.cn> (accessed on 7 December 2019). [\[CrossRef\]](#)
31. Xu, X. China GDP Spatial Distribution Kilometer Grid Dataset. Data Registration and Publishing System of Resource and Environmental Science Data Center of the Chinese Academy of Sciences. 2017. Available online: <http://www.resdc.cn> (accessed on 7 December 2019). [\[CrossRef\]](#)
32. Jia, C.; Du, Y.; Wang, S.; Bai, T.; Fei, T. Measuring the vibrancy of urban neighborhoods using mobile phone data with an improved PageRank algorithm. *Trans. GIS* **2019**, *23*, 241–258. [\[CrossRef\]](#)
33. Brin, S.; Page, L. The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.* **1998**, *30*, 107–117. [\[CrossRef\]](#)
34. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [\[CrossRef\]](#)
35. Sung, H.; Lee, S. Residential built environment and walking activity: Empirical evidence of Jane Jacobs' urban vitality. *Transp. Res. Part D Transp. Environ.* **2015**, *41*, 318–329. [\[CrossRef\]](#)
36. Zhang, A.; Li, W.; Wu, J.; Lin, J.; Chu, J.; Xia, C. How can the urban landscape affect urban vitality at the street block level? A case study of 15 metropolises in China. *Environ. Plan. B Urban Anal. City Sci.* **2020**. [\[CrossRef\]](#)
37. Lu, S.; Shi, C.; Yang, X. Impacts of built environment on urban vitality: Regression analyses of Beijing and Chengdu, China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4592. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Landry, C. Urban vitality: A new source of urban competitiveness. *Archis* **2000**, *12*, 8–13.
39. Liu, L.; Xu, Y.-L.; Jiang, S.-H.; Wu, Q.-M. Evaluation of Urban Vitality Based on Fuzzy Matter-Element Model. *Geogr. Geo-Inf. Sci.* **2010**, *26*, 73–77.
40. Lan, F.; Gong, X.; Da, H.; Wen, H. How do population inflow and social infrastructure affect urban vitality? Evidence from 35 large-and medium-sized cities in China. *Cities* **2020**, *100*, 102454. [\[CrossRef\]](#)
41. Liu, M.; Hu, Y.-M.; Li, C.-L. Landscape metrics for three-dimensional urban building pattern recognition. *Appl. Geogr.* **2017**, *87*, 66–72. [\[CrossRef\]](#)
42. Mouratidis, K.; Poortinga, W. Built environment, urban vitality and social cohesion: Do vibrant neighborhoods foster strong communities? *Landsc. Urban Plan.* **2020**, *204*, 103951. [\[CrossRef\]](#)